

TAI, SHUICHI

19980203

Feb. 3, 1998

PAT-NO: JP410031551A

DOCUMENT-IDENTIFIER: JP 10031551 A

TITLE: HUMAN INTERFACE SYSTEM AND HIGH-SPEED
MOVING BODY
POSITION DETECTING DEVICE USING THE SAME

----- KWIC -----

Abstract Text - FPAR (1):

PROBLEM TO BE SOLVED: To provide a more compact human interface system which detects and recognizes a hand gesture and a body action by an image sensor, has a wide application range, and operates fast.

Abstract Text - FPAR (2):

SOLUTION: A human's gesture is inputted as image information by an image sensor 12 consisting of an artificial retina chip and a CCD and a human voice is inputted as speech information by a speech input device 13 consisting of microphone, etc.; and the gesture and voice are recognized by a microcomputer 14 through incorporated gesture and voice recognizing algorithm 16 and then a control signal is supplied to a controlled device 17 consisting of multimedia equipment, a home appliance, a game machine, care device, etc., as a trailing state.

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-31551

(43)公開日 平成10年(1998)2月3日

| (51)Int.Cl.* | 識別記号 | 序内整理番号 | F I | 技術表示箇所 |
|----------------|-------|--------|--------------|---------|
| G 06 F 3/033 | 3 1 0 | | G 06 F 3/033 | 3 1 0 Y |
| G 06 T 1/00 | | | G 10 L 3/00 | 5 7 1 G |
| G 10 L 3/00 | 5 7 1 | | G 06 F 3/16 | 3 2 0 A |
| // G 06 F 3/16 | 3 2 0 | | 15/62 | 3 8 0 |

審査請求 未請求 請求項の数 8 OL (全 8 頁)

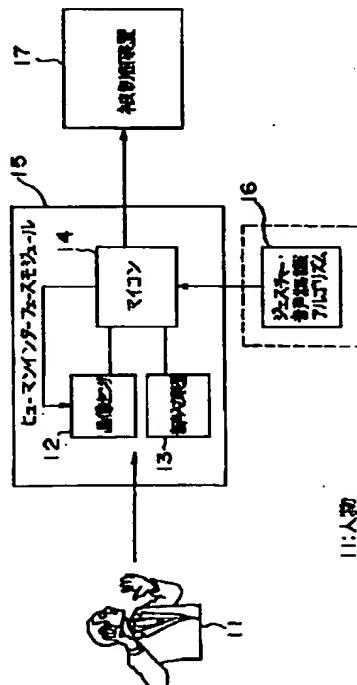
| | |
|----------------------------|---|
| (21)出願番号 特願平8-184951 | (71)出願人 三菱電機株式会社 東京都千代田区丸の内二丁目2番3号 |
| (22)出願日 平成8年(1996)7月15日 | (72)発明者 田井 修市 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社内 |
| | (72)発明者 久間 和生 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社内 |
| | (72)発明者 田中 健一 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社内 |
| | (74)代理人 弁理士 曾我 道照 (外6名) 最終頁に続く |

(54)【発明の名称】 ヒューマンインターフェースシステムおよびこれを使用した高速移動物体位置検出装置

(57)【要約】

【課題】 画像センサによってハンドジェスチャーやボディアクションを検出・認識する、適用範囲が広く、高速で動作し、さらにコンパクトなヒューマンインターフェースシステムを提供する。

【解決手段】 人工網膜チップやCCDからなる画像センサ12により人間のジェスチャーを画像情報として入力し、さらにマイク等からなる音声入力装置13により人間の音声を音声情報として入力し、内蔵されたジェスチャー・音声認識アルゴリズム16に基づいてジェスチャーおよび音声をマイコン14で認識処理し、これに基づいて制御信号を後段のマルチメディア機器、家電製品、ゲーム機あるいは介護装置等からなる被制御装置17に供給する。



【特許請求の範囲】

【請求項1】 人間の体による表現を入力として取り込むためのヒューマンインターフェースシステムであつて、

人間のジェスチャーを画像情報として入力するジェスチャー入力手段と、
ジェスチャー認識のためのアルゴリズムが格納されているアルゴリズム格納手段と、
このアルゴリズム格納手段のアルゴリズムに基づいて上記ジェスチャー入力手段で入力された画像情報を認識処理し、これに基づく制御信号を後段に供給する情報認識処理手段と、
を備えたヒューマンインターフェースシステム。

【請求項2】 人間の音声を音声情報として入力する音声入力手段をさらに備え、

上記アルゴリズム格納手段が、音声認識のためのアルゴリズムをさらに格納し、上記情報認識処理手段が、上記アルゴリズム格納手段のアルゴリズムに基づいて、入力された画像情報および音声情報を認識処理し、これらを統合した結果に基づく制御信号を後段に供給することを特徴とする請求項1に記載のヒューマンインターフェースシステム。

【請求項3】 上記情報認識処理手段がハンドジェスチャー認識を行い、上記ハンドジェスチャー認識は手の形状、位置、傾きによって行われ、認識結果は手を長方形で近似して、その重心のX-Y座標、サイズ、傾き角などで定量的に検出し、種々のハンドジェスチャーを認識することを特徴とする請求項1または2に記載のヒューマンインターフェースシステム。

【請求項4】 上記情報認識処理手段がボディアクション認識を行い、上記ボディアクション認識は人間の腕の位置および身体の形状の認識、あるいはこれらの予め格納されている基本ボディアクションとの比較により、与えられたボディアクションを認識することを特徴とする請求項1ないし3のいずれかに記載のヒューマンインターフェースシステム。

【請求項5】 遠隔制御のための入力機構として設けられたことを特徴とする請求項1ないし4のいずれかに記載のヒューマンインターフェースシステム。

【請求項6】 上記ジェスチャー入力手段が前処理機能を有する人工網膜チップからなる請求項1ないし5のいずれかに記載のヒューマンインターフェースシステム。

【請求項7】 上記ジェスチャー入力手段がCCDからなる請求項1ないし5のいずれかに記載のヒューマンインターフェースシステム。

【請求項8】 移動物体のX-Y座標検出を行うパターンマッチング機能を有する人工網膜チップを搭載した第1の入力手段と、

同時刻の上記移動物体のZ座標検出を行うパターンマッチング機能を有する人工網膜チップを搭載した第2の入

力手段と、

上記第1および第2の入力手段で得られたデータから上記移動物体の軌跡を計算する制御処理手段と、
からなる高速移動物体位置検出装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明はコンピュータと人間とがインタラクティブに作用するための知的で柔軟なヒューマンインターフェースシステムに関するものである。

【0002】

【従来の技術】従来のヒューマンインターフェースシステムの一例として、ハンドジェスチャー認識システムについて説明する。図8は例えば、電子情報通信学会論文誌、D-II、Vol. J73-D-II、No. 12、PP-1985~1992(1990年)に掲載された、ハンドジェスチャー認識システムの構成図である。図8において、1は手形状入力装置(市販のデータグラブTM "Data Glove TM" を使用)である。データグラブTMは米国VPL社の製品で、指の曲げ情報を光ファイバ内の光量で測定し、親指の第2、3関節、他の指の第1、2関節の10関節について個人差を補正した曲げ角度を測定する。磁気発生器と磁気センサを用いた米国マクダネルダグラス(MacDonnell Douglas)社の3次元デジタイザ(3SPACE TRACKER TM)を組合せることで、手の向きと手の3次元位置も測定できる。

【0003】2は補正器、3はマイク等からなる音声認識装置、4はスピーカ等からなる音声合成装置、5は画像表示装置、6はワークステーション、7は手の形・動きを認識するための認識プログラム、8は音声認識結果を解析するための解析プログラム、9は手の形・動きの認識結果と音声解析結果を統合し、指示された内容を音声または画像の形で出力するための変換プログラム、10は画像データベースである。

【0004】次に動作について説明する。ハンドジェスチャーは手形状入力装置1と個人差の補正器2によってワークステーション6に入力され、手の形・動きを認識するための認識プログラム7によって認識される。また、音声情報は音声認識装置3によって認識された後、

ワークステーション6の音声認識結果を解析するための解析プログラム8に送られる。

【0005】ハンドジェスチャーと音声の認識結果は、手の形・動きの認識結果と音声認識結果を統合し、指示された内容を音声または画像の形で出力するための変換プログラム9によって、外部の音声合成装置4あるいは画像表示装置5に送られる。

【0006】

【発明が解決しようとする課題】従来のヒューマンインターフェースシステムは以上のように構成されているので、ハンドジェスチャー認識のために、手袋のような手

形状入力装置(図8の1参照)が必要であり、情報処理にワークステーションが必要であるなど、非常に繁雑で大がかりなシステムであった。従って、汎用的なヒューマンインターフェースシステムとしての使用は困難である等の問題があった。

【0007】この発明は、上記のような課題を解決するためになされたもので、画像センサによってハンドジェスチャーやボディアクションを検出・認識する、適用範囲が広く、高速で動作し、さらにコンパクトなヒューマンインターフェースおよびこれを使用した高速移動物体位置検出装置を提供することを目的とする。

【0008】

【課題を解決するための手段】上記の目的に鑑み、この発明の第1の発明は、人間の体による表現を入力として取り込むためのヒューマンインターフェースシステムであって、人間のジェスチャーを画像情報として入力するジェスチャー入力手段と、ジェスチャー認識のためのアルゴリズムが格納されているアルゴリズム格納手段と、このアルゴリズム格納手段のアルゴリズムに基づいて上記ジェスチャー入力手段で入力された画像情報を認識処理し、これに基づく制御信号を後段に供給する情報認識処理手段と、を備えたヒューマンインターフェースシステムにある。

【0009】この発明の第2の発明は、人間の音声を音声情報として入力する音声入力手段をさらに備え、上記アルゴリズム格納手段が、音声認識のためのアルゴリズムをさらに格納し、上記情報認識処理手段が、上記アルゴリズム格納手段のアルゴリズムに基づいて、入力された画像情報および音声情報を認識処理し、これらを統合した結果に基づく制御信号を後段に供給することを特徴とする請求項1に記載のヒューマンインターフェースシステムにある。

【0010】この発明の第3の発明は、上記情報認識処理手段がハンドジェスチャー認識を行い、上記ハンドジェスチャー認識は手の形状、位置、傾きによって行われ、認識結果は手を長方形で近似して、その重心のX-Y座標、サイズ、傾き角などで定量的に検出し、種々のハンドジェスチャーを認識することを特徴とする請求項1または2に記載のヒューマンインターフェースシステムにある。

【0011】この発明の第4の発明は、上記情報認識処理手段がボディアクション認識を行い、上記ボディアクション認識は人間の腕の位置および身体の形状の認識、あるいはこれらの予め格納されている基本ボディアクションとの比較により、与えられたボディアクションを認識することを特徴とする請求項1ないし3のいずれかに記載のヒューマンインターフェースシステムにある。

【0012】この発明の第5の発明は、遠隔制御のための入力機構として設けられたことを特徴とする請求項1ないし4のいずれかに記載のヒューマンインターフェー

スシステムにある。

【0013】この発明の第6の発明は、上記ジェスチャー入力手段が前処理機能を有する人工網膜チップからなる請求項1ないし5のいずれかに記載のヒューマンインターフェースシステムにある。

【0014】この発明の第7の発明は、上記ジェスチャー入力手段がCCDからなる請求項1ないし5のいずれかに記載のヒューマンインターフェースシステムにある。

10 【0015】この発明の第8の発明は、移動物体のX-Y座標検出を行うパターンマッチング機能を有する人工網膜チップを搭載した第1の入力手段と、同時刻の上記移動物体のZ座標検出を行うパターンマッチング機能を有する人工網膜チップを搭載した第2の入力手段と、上記第1および第2の入力手段で得られたデータから上記移動物体の軌跡を計算する制御処理手段と、からなる高速移動物体位置検出装置にある。

【0016】

【発明の実施の形態】以下、この発明を各実施の形態に従って説明する。

実施の形態1. 図1はこの発明の一実施の形態によるヒューマンインターフェースシステムの構成を示すものである。図において11は人物、12は人工網膜チップおよびCCD等からなる画像センサ、13はマイク等からなる音声入力装置、14はマイクロコンピュータ(以下マイコン)、15はヒューマンインターフェースモジュール、16はジェスチャー認識および音声認識アルゴリズム(ソフトウェア)、17はパソコンを含むマルチメディア機器、家電製品(テレビ、エアコン、オーディオなど)、ゲーム機、介護装置などからなる被制御装置である。

【0017】なお、画像センサ12が人間のジェスチャーを画像情報として入力するジェスチャー入力手段を構成し、音声入力装置13が人間の音声を音声情報として入力する音声入力手段を構成し、ジェスチャー・音声認識アルゴリズム16がアルゴリズム格納手段を構成し、マイコン14が情報認識処理手段を構成する。

【0018】次に動作について説明する。人物11はハンドジェスチャーやボディアクションによって自分の意思を表現する。画像センサ12には人工網膜チップあるいはCCDを使用する。人工網膜チップは図2に示す構造を有している。ここで18は人工網膜チップ基板、19はランダムアクセススキャナ、20はマルチブレクサ、21は画素コア回路、22は感度制御信号、23は入力画像、24は出力画像である。

【0019】画素コア回路21は光検出感度を正から負まで変えることができるものであり、ランダムアクセススキャナ19は画素コア回路21の光検出感度を変えるための制御信号22を発生するものである。例えば、画素コア回路21の1行目と2行目の感度をそれぞれ1、

-1とし、3行目～最終行の感度を0(光が当たっても電流が流れないとすれば、1行目と2行目では照射された入力画像23の差信号が得られる。従って、この感度制御信号22のペア(1, -1)を1行目から最終行までスキャンすることによって、入力画像23のエッジ検出画像が出力画像24として得られる。

【0020】感度制御信号22を1, 0, 0, …とすれば、入力画像23をそのままの形で検出することができる(CCDの機能)。また、感度制御信号22を全て1とすれば、入力画像23の射影信号が出力画像24に得られる。人工網膜チップは画像検出・処理が高速(フレームレート: 1 kHz以上)に行えるという特長も有している。人工網膜チップで前処理された信号は、マイコン14によって処理され、人物のハンドジェスチャーやボディアクションの認識処理が行われる。なお、マイコン14は上述の人工網膜チップでの画像処理内容の制御も行う。

【0021】音声はマイクロホンなどの音声入力装置13によって入力され、マイコン14などの信号処理装置によってスペクトル分解、音韻分解などが行われ、認識される。

【0022】これらのハンドジェスチャーやボディアクション認識と音声認識の結果は統合されて、パソコン、家電製品(テレビ、エアコン、オーディオなど)、ゲーム機、介護装置などの被制御装置17に制御信号として入力され、人物11の意思のとおりに被制御装置17を動作させることができる。これらの認識アルゴリズム16は、例えば、情報処理装置であるマイコン14などに付随して搭載されている。

【0023】なお、上記実施の形態では画像センサとして、画像の前処理機能(エッジ検出、射影検出など)を有する人工網膜チップを使用したが、画像処理機能を有しないCCDを使用することも可能である。CCDを使用した場合には、画像の前処理と信号処理を共にマイコン14等の情報処理装置で行う必要がある。

【0024】さらにこの発明では、人間の音声を入力するための音声入力装置13は必要不可欠なものではなく、なくてもよいことは言うまでもない。

【0025】実施の形態2. 図3は1つの実施の形態である、ハンドジェスチャーによる情報入力の模様を示したものである。図3において、15aは人工網膜チップとマイコンを搭載したヒューマンインターフェースモジュール、17はパソコン等からなる被制御装置、11aは使用者である。図3に示すように、使用者11aのハンドジェスチャーをコンパクトなヒューマンインターフェースモジュール15aによって認識し、このハンドジェスチャーに基づいてパソコンや家電製品、ゲーム機、介護装置などの被制御装置17を制御できる。

【0026】図4の(a)～(d)は、ハンドジェスチャー認識の例である。図において、28は人工網膜チップの

撮像領域(モニタ画面)、29は手、30は手の形を長方形で近似したもの、31は重心である。ハンドジェスチャーは図4のように、手の形を長方形で近似して、その重心座標(X1, Y1)、サイズ(d1 × d2)、傾き角(θ)によって認識される。このように、単純かつ少量の情報によって、さまざまなハンドジェスチャー認識が行える。

【0027】実施の形態3. 図5は別の1つの実施の形態の特徴である、人間のボディアクションを認識することによって、使用者の意思(スイッチオン、オフ等)を認識するものである。ボディアクションの例を図5の(a)～(f)に示す。図において、32a～32fはボディアクションの特に腕の位置(上下、左右、斜め、水平等)を示している。これら以外にも、身体の左右の傾き、屈み込み、ジャンプ等の身体の形状のボディアクションの認識も可能である。また、予め登録された基本ボディアクションと入力されたボディアクションとを比較することによって、ボディアクションの認識を行うこともできる。

【0028】実施の形態4. 図6は別の1つの実施の形態の特徴である、ハンドジェスチャーやボディアクションを入力としたゲーム機への応用の一例である。図において、15bはヒューマンインターフェースモジュール、33はゲーム画面、11bはプレーヤー、35はゲームの登場人物である。図に示すように、プレーヤー11bのボディアクションのとおりに、ゲームの登場人物35が動作する。図6では、フライングゲームに適用した例を示している。プレーヤー11bの動作どおりに、登場人物35が空を飛ぶため、プレーヤー11bはあたかも自分が空を飛んでいるような気分が味わえる。

【0029】このようにハンドジェスチャーやボディアクションを入力として、ジョイステックの不要なゲーム機や、登場人物がプレーヤーと同じ動作を行うゲーム機、あるいはキーボード、マウス、スイッチ、ダイヤル、リモコンなどを用いることなく、コンピュータ、家電製品、介護装置を制御できるヒューマンインターフェースシステムが実現できる。

【0030】実施の形態5. 次に、図7には別の1つの実施の形態のヒューマンインターフェースシステムを使用した高速移動物体位置検出装置の構成を示す。図7では高速で移動する物体の位置検出を行うものである。図において、36a、36bは人工網膜チップ搭載カメラ、37a～37eは一定時間間隔を空けて表示した移動物体の存在する平面、38a～38eは移動物体、39は人工網膜チップ搭載カメラ36a、36bからの入力に基づいて移動物体の軌跡を計算処理するマイコンである。

【0031】なお、人工網膜チップ搭載カメラ36a、36bが第1および第2の入力手段、マイコン39が制御処理手段を構成する。

【0032】次に動作について説明する。人工網膜チップはパターンマッチング機能を有している。すなわち、入力画像中から特定パターンの画像を抽出することができるため、移動物体の形状(ボールならば円形)をパターンマッチングすることにより、移動物体のみを検出できる。従って、即座にその位置(座標)の検出が行える。人工網膜チップ搭載カメラ36aは移動物体のX-Y座標を検出するためのもの、人工網膜チップ搭載カメラ36bは移動物体のZ座標を検出するためのものである。

【0033】すなわち、時間t1において平面37a上の移動物体38aのX-Y座標を人工網膜チップ搭載カメラ36aで、Z座標を人工網膜チップ搭載カメラ36bで検出することにより、時間t1における移動物体38aの座標が分かる。同様に、時間t2、t3、…においても移動物体38b、38c、…の座標をそれぞれ検出することができる。人工網膜チップは高速画像検出・処理が行える(フレームレート: 1 kHz)ため、1 msec毎に移動物体の座標検出が可能である。

【0034】移動物体を、例えば投げられたボールとすれば、その軌跡は方物線(二次曲線 $y = aX^2 + bX + c$)を描くため、最低異なる3つの時間のX-Y座標とZ座標(距離を示す)がわかれば、マイコン39により軌跡は計算によって求めることができる。軌跡が分かれればボールの落下地点も分かる。

【0035】

【発明の効果】以上のようにこの発明の第1の発明では、人間のジェスチャーを画像情報として入力し、内蔵されたアルゴリズムによりジェスチャーを認識処理し、これに基づいて制御信号を後段に供給するようにしたので、マルチメディア機器、家電製品(テレビ、エアコン、オーディオ等)、ゲーム機および介護装置等へ人間のジェスチャーにより情報を入力できるヒューマンインターフェースシステムを提供できる等の効果が得られる。

【0036】この発明の第2の発明では、さらに人間の音声を音声情報として入力し、内蔵された音声認識のためのアルゴリズムにより音声を認識処理し、画像情報認識処理と統合した結果に基づく制御信号を後段に供給するようにしたので、さらに複雑で高度な情報を入力できるヒューマンインターフェースシステムを提供できる等の効果が得られる。

【0037】この発明の第3の発明では、ハンドジェスチャー認識を行い、ハンドジェスチャー認識は手の形状、位置、傾きによって行われ、認識結果は手を長方形で近似して、その重心のX-Y座標、サイズ、傾き角などで定量的に検出し、種々のハンドジェスチャーを認識するようにしたので、ハンドジェスチャーという極めて容易な動作で情報入力が行えるヒューマンインターフェースシステムを提供できる等の効果が得られる。

【0038】この発明の第4の発明では、ボディアクシ

10

ョン認識を行い、ボディアクション認識は人間の腕の位置および身体の形状の認識、あるいはこれらの予め格納されている基本ボディアクションとの比較により、与えられたボディアクションを認識するようにしたので、より認識の容易な人間の体による表現に基づく情報入力が行え、さらにハンドジェスチャーと組合わせれば、より複雑で高度な情報入力が行えるヒューマンインターフェースシステムを提供できる等の効果が得られる。

【0039】この発明の第5の発明では、人間のハンド

ジェスチャーやボディアクションおよび音声によって、マルチメディア機器、家電製品(テレビ、エアコン、オーディオ等)、ゲーム機、介護装置等への情報を遠隔入力できるため、キーボード、マウス、ジョイスティック、ボタン、ダイヤル、リモコンなどが不要のインターフェースを実現でき、老人、子供、病人、身体障害者等のための人に優しいヒューマンインターフェースシステムを提供できる等の効果が得られる。

【0040】この発明の第6の発明では、ジェスチャーを画像情報として入力する入力手段を前処理機能を有する人工網膜チップとしたので、高速で動作し、さらに従来の手形状入力装置を使用したものに比べてコンパクトなヒューマンインターフェースシステムを提供できる等の効果が得られる。

【0041】この発明の第7の発明では、ジェスチャーを画像情報として入力する入力手段をCCDとしたので、入手が容易で、かつ従来の手形状入力装置を使用したものに比べてコンパクトなヒューマンインターフェースシステムを提供できる等の効果が得られる。

【0042】この発明の第8の発明では、移動物体のX-Y座標検出を行うパターンマッチング機能を有する人工網膜チップを搭載した第1の入力手段と、同時刻の移動物体のZ座標検出を行うパターンマッチング機能を有する人工網膜チップを搭載した第2の入力手段とで得られたデータから移動物体の軌跡を計算するようにしたので、移動物体の位置を高速で検出できる高速移動物体位置検出装置を提供できる等の効果が得られる。.

【図面の簡単な説明】

【図1】 この発明の一実施の形態によるヒューマンインターフェースシステムの構成図である。

【図2】 人工網膜チップの構成図である。

【図3】 この発明の1つの実施の形態の特徴であるハンドジェスチャーによる情報入力の模様を示した図である。

【図4】 (a)～(d)はハンドジェスチャー認識の例を示す図である。

【図5】 この発明の1つの実施の形態の特徴である人間のボディアクションを認識する場合のボディアクションの例を示す図である。

【図6】 この発明の1つの実施の形態の特徴であるハンドジェスチャーとボディアクションを入力としたゲー

50

ム機への応用の一例を示す図である。

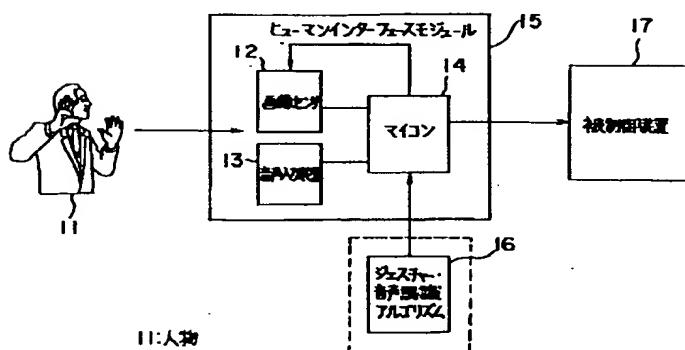
【図7】 この発明の別の実施の形態によるヒューマンインターフェースシステムを使用した高速移動物体位置検出装置の構成を示す図である。

【図8】 従来のハンドジェスチャー認識システムの構成図である。

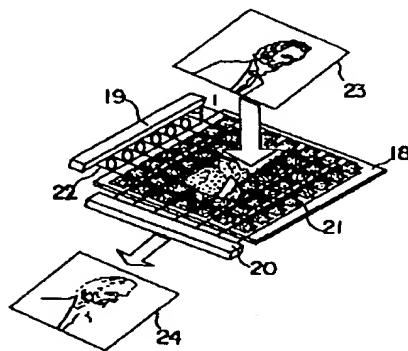
【符号の説明】

11 人物、11a 使用者、11b プレーヤー、12 画像センサ、13 音声入力装置、14, 39 マイコン、15, 15a, 15b ヒューマンインターフェースモジュール、16 ジェスチャー・音声認識アルゴリズム、17 被制御装置、18 人工網膜チップ基板、33 ゲーム画面、35 登場人物、36a, 36b 人工網膜チップ搭載カメラ。

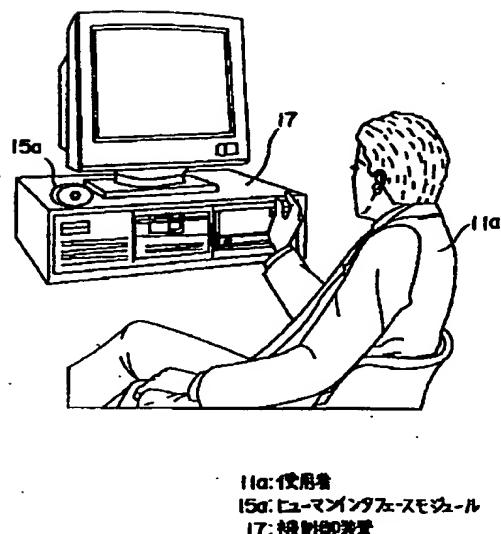
【図1】



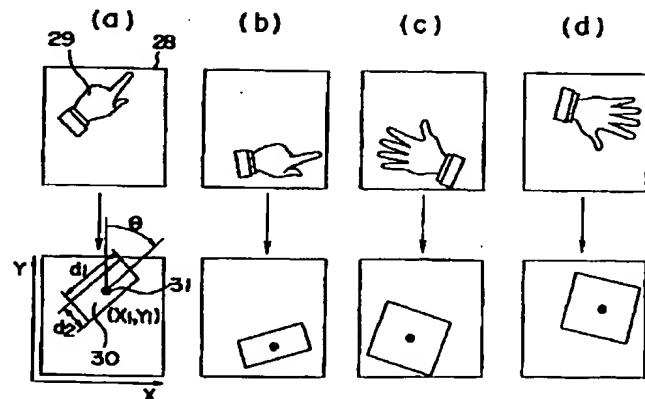
【図2】



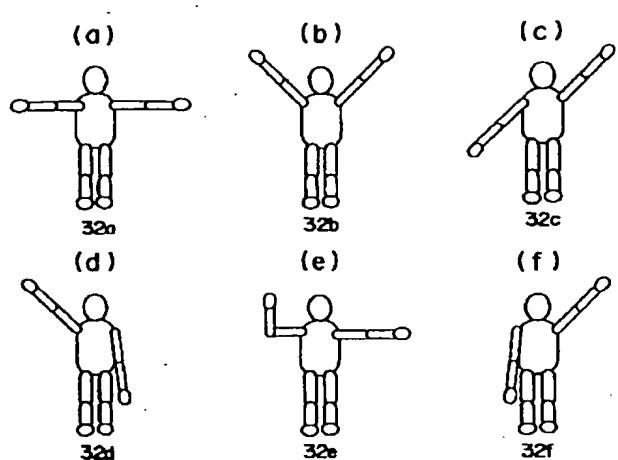
【図3】



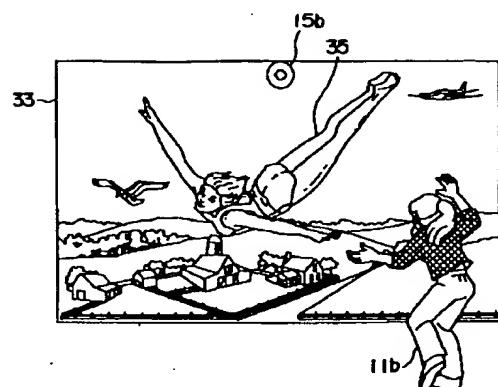
【図4】



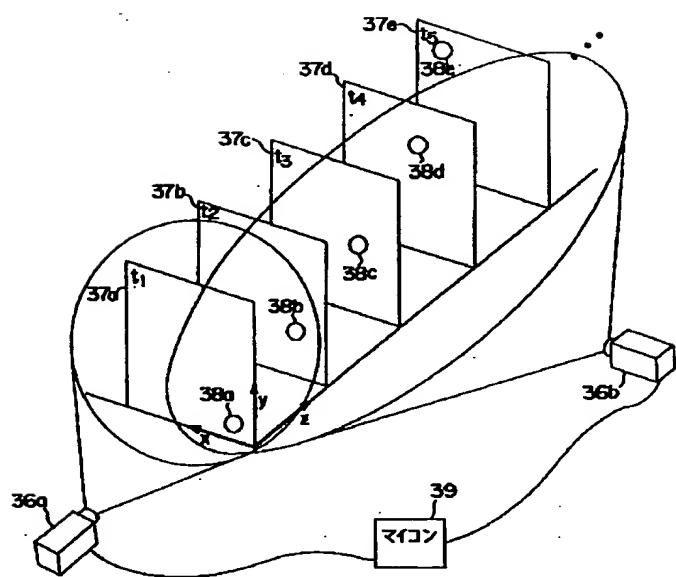
【図5】



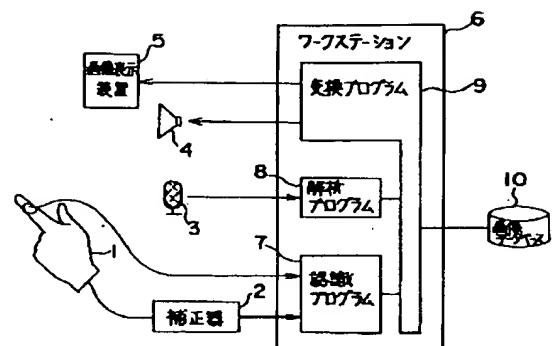
【図6】



【図7】



【図8】



フロントページの続き

(72)発明者 太田 淳

東京都千代田区丸の内二丁目2番3号 三
菱電機株式会社内

(72)発明者 小守 伸史

東京都千代田区丸の内二丁目2番3号 三
菱電機株式会社内

(72)発明者 ウィリアム・ティ・フリーマン

アメリカ合衆国、マサチューセッツ州、ニ
ュートン、アラートンロード 174